

09/24/99

**Match & Return**

ASSISTANT COMMISSIONER FOR PATENTS  
Washington, D.C. 20231

Sir:

Transmitted herewith for filing under 37 C.F.R. §1.53(b) is the patent application of:  
Inventor(s): Gertrud HÖTTEN, Helge NEIDHARDT, Rolf BECHTOLD and Jen POHL

For: **GROWTH/DIFFERENTIATION FACTORS OF THE TGF-β FAMILY**

This application is a divisional of Application No. 08/289,222

☒ Return Receipt Postcard

☒ Specification (34 pages)

☒ -3- sheets of drawings

☒ A copy of Sequence Listing and Statement with disk

☒ A Preliminary Amendment

☒ Declaration and Power of Attorney

☒ Copy from a prior application

☒ The disclosure of the prior application, from which a copy of the declaration is supplied as noted above is considered as being a part of the disclosure of the accompanying application and is hereby incorporated by reference therein.

☒ Assignment was filed in parent Appln. No. 08/289,222, Reel/Frame 7426/0515.

☒ Priority of German application, Serial No. P44 23 190.3 filed 01/07/1994 and European application, Serial No. 92 102 324.8 filed 12/02/1992 is claimed under 35 U.S.C. §119.

☒ An Information Disclosure Statement with PTO-1449

☒ A filing fee, calculated as shown below, including claims added or deleted in the Preliminary Amendment (**Check No. 20576**):

(Col. 1)

(Col. 2)

FOR:	No. Filed	No. Extra
BASIC FEE		
TOTAL CLAIMS	12 - 20 =	-0-
INDEP CLAIMS	6 - 3 =	3
_ MULTIPLE DEPENDENT CLAIM PRESENTED		

Small Entity

RATE	FEE
	\$380
× 9 =	
× 39 =	
+130 =	
TOTAL	

Other Than A Small Entity

RATE	FEE
	\$760
× 18 =	
× 78 =	\$234
+260 =	
	\$994

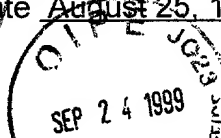
Respectfully submitted,  
NIKAIDO, MARMELESTEIN, MURRAY & ORAM LLP

By: Monica Chin Kitts  
Reg. No. 36,105

Metropolitan Square  
655 Fifteenth Street, N.W.  
Suite 330 - G Street Lobby  
Washington, D.C. 20005-5701  
(202) 638-5000  
MCK:ecm

Case Docket No. P564-9021

Date August 25, 1999



**RECEIVED**

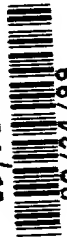
SEP 24 1999

1999

Publishing Division

JC474 U.S. PTO

09/901556



09/24/99

RECEIVED

FEB - 9 2000

RECEIVED

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the application of:

HOTTEN et al.

Serial Number: Unknown

Filed: August 25, 1999

For: GROWTH/DIFFERENTIATION FACTORS OF THE TGF- $\beta$  FAMILY

Group Art Unit: Unknown

Examiner: Unknown

**STATEMENT UNDER 37 CFR §1.821(C)**

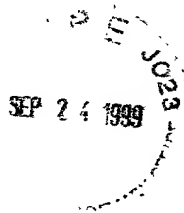
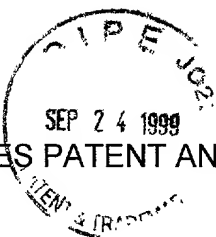
Assistant Commissioner  
of Patents and Trademarks  
Washington, D.C. 20231

August 25, 1999

Sir:

In accordance with 37 C.F.R.1.821(C), applicants are submitting herewith the Sequence Listing for the above-identified application both in paper copy form and in computer readable form.

The name of the file on the computer readable form is 5649021.APP. The paper copy and the computer readable form are the same.



#18/M  
w/cont  
TDP  
N.E



654260-95470660

In the event that this paper is not considered to be timely filed, applicants hereby petition for an appropriate extension of time. The fee for any such extension may be charged to our Deposit Account No. 14-1060, along with any other fees with respect to this paper.

Respectfully submitted,  
NIKAIDO, MARMELESTEIN, MURRAY & ORAM LLP

  
Monica Chin Kitts  
Attorney for Applicants  
Registration No. 36,105

Atty. Docket No. P564-9021  
Metropolitan Square  
655 Fifteenth Street, N.W.  
Suite 330 - G Street Lobby  
Washington, D.C. 20005-5701  
(202) 638-5000

MCK/TPC

M 149  
E 1-7-00  
N.E



Serial No.: unknown

For: GROWTH/DIFFERENTIATION FACTORS OF THE TGF- $\beta$  FAMILY

Assistant Commissioner for Patents  
Washington, D.C. 20231

August 25, 1999

Prior to calculation of the filing fee and prior to the examination of this application, please amend the above-identified application as follows:

Kindly cancel claims 1-19 without prejudice or disclaimer.

Please add the following new claims to the application.

--20. An antibody or antibody fragment which specifically binds to a protein of the

TGF- $\beta$  family wherein said protein is encoded by a DNA comprising a nucleotide sequence selected from the following group:

- (a) the nucleotide sequence as shown in SEQ ID NO:1,
- (b) a nucleotide sequence which is degenerate as a result of the genetic code to the nucleotide sequence of (a), and
- (c) fragments of (a) or (b) which encode a protein which has essentially the same

cartilage or bone inducing activities as a mature protein encoded by the nucleotide sequence of SEQ ID NO:1.

21. The antibody according to claim 20, wherein said antibody is a monoclonal antibody.

22. An antibody or antibody fragment according to claim 20, which specifically binds to a protein of the TGF- $\beta$  family wherein said protein comprises the amino acid sequence according to SEQ ID NO:3.

23. The antibody according to claim 22, wherein said antibody is a monoclonal antibody.

24. An antibody or antibody fragment which specifically binds a protein of the TGF- $\beta$  family, wherein said protein is encoded by a DNA comprising a nucleotide sequence selected from the following group:

(a) the nucleotide sequence as shown in SEQ ID NO:2,

(b) a nucleotide sequence which is degenerate as a result of the genetic code to the DNA of (a),

(c) a nucleotide sequence which hybridizes under the following stringent hybridization conditions to the DNA in (a), or (b): hybridization at a salt concentration of 4X SSC at 62°-66°C followed by a one-hour wash with 0.1X SSC and 0.1% SDS at 62°-66°C, and

(d) fragments of (a), (b) or (c) which encode a protein which has essentially the same cartilage or bone inducing activity as a mature protein encoded by the nucleotide sequence of SEQ ID NO:2.

25. An antibody or antibody fragment according to claim 24, wherein said protein comprises the amino acid sequence according to SEQ ID NO:4.

26. The antibody according to claim 25, wherein said antibody is a monoclonal antibody.

27. The antibody according to claim 24, wherein said antibody is a monoclonal antibody.

28. A method for detecting a protein of the TGF- $\beta$  family,  
comprising incubating an antibody or antibody fragment which specifically binds to a protein of the TGF- $\beta$  family with a sample suspected of containing said protein, and detecting any antibody/protein complex formed as an indication of the presence of said protein,

wherein said protein is encoded by a DNA comprising a nucleotide sequence selected from the following group:

(a) the nucleotide sequence as shown in SEQ ID NO:1,

(b) a nucleotide sequence which is degenerate as a result of the genetic code to the nucleotide sequence of (a), and

(c) fragments of (a) or (b) which encode a protein which has essentially the same cartilage or bone inducing activities as a mature protein encoded by the nucleotide sequence of SEQ ID NO:1.

29. A method for detecting a protein of the TGF- $\beta$  family, comprising  
incubating an antibody or antibody fragment which specifically binds to said protein of the TGF- $\beta$  family with a sample suspected of containing said protein, and  
detecting any antibody/protein complex formed as an indication of the presence of said protein,

wherein said protein is encoded by a DNA comprising a nucleotide sequence selected from the following group:

(a) the nucleotide sequence as shown in SEQ ID NO:2,  
(b) a nucleotide sequence which is degenerate as a result of the genetic code to the DNA of (a),

(c) a nucleotide sequence which hybridizes under the following stringent hybridization conditions to the DNA in (a), or (b): hybridization at a salt concentration of 4X SSC at 62°-66°C followed by a one-hour wash with 0.1X SSC and 0.1% SDS at 62°-66°C, and

(d) fragments of (a), (b) or (c) which encode a protein which has essentially the same cartilage or bone inducing activity as a mature protein encoded by the nucleotide sequence of SEQ ID NO:2.

30. A kit for detecting a protein of the TGF- $\beta$  family, comprising

an antibody or antibody fragment which specifically binds to a protein of the TGF- $\beta$  family, and

a reaction buffer,

wherein said protein is encoded by a DNA comprising a nucleotide sequence selected from the following group:

(a) the nucleotide sequence as shown in SEQ ID NO:1,

(b) a nucleotide sequence which is degenerate as a result of the genetic code to the nucleotide sequence of (a), and

(c) fragments of (a) or (b) which encode a protein which has essentially the same cartilage or bone inducing activities as a mature protein encoded by the nucleotide sequence of SEQ ID NO:1.

31. A kit for detecting a protein of the TGF- $\beta$  family, comprising

an antibody or antibody fragment which specifically binds to a protein of the TGF- $\beta$  family, and

a reaction buffer,

wherein said protein is encoded by a DNA comprising a nucleotide sequence selected from the following group:

(a) the nucleotide sequence as shown in SEQ ID NO:2,

(b) a nucleotide sequence which is degenerate as a result of the genetic code to the DNA of (a),

(c) a nucleotide sequence which hybridizes under the following stringent hybridization conditions to the DNA in (a), or (b): hybridization at a salt concentration of



4X SSC at 62°-66°C followed by a one-hour wash with 0.1X SSC and 0.1% SDS at 62°-66°C, and

(d) fragments of (a), (b) or (c) which encode a protein which has essentially the same cartilage or bone inducing activity as a mature protein encoded by the nucleotide sequence of SEQ ID NO:2. --

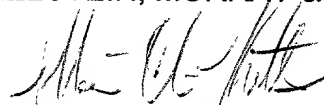
#### REMARKS

The above amendments have been made to put the application into better condition for examination.

In the event that any fees are due in connection with this paper, please charge our Deposit Account No. 14-1060.

Respectfully submitted,

NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP

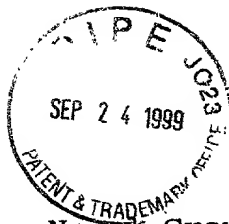


Monica Chin Kitts  
Attorney for Applicants  
Registration No. 36,105

Atty. Docket No. 564-9021  
Metropolitan Square  
655 Fifteenth Street, N.W.  
Suite 330 - G Street Lobby  
Washington, D.C. 20005-5701  
(202) 638-5000

MCK/mck

654260-9310650



DNA Sequences Encoding Novel Growth/  
Differentiation Factors

09901556 "092499

The present invention relates to DNA sequences encoding novel growth/differentiation factors of the TGF- $\beta$  family. In particular, it relates to novel DNA sequences encoding TGF- $\beta$ -like proteins, to the isolation of said DNA sequences, to expression plasmids containing said DNA, to microorganisms transformed by said expression plasmid, to the production of said protein by culturing said transformant, and to pharmaceutical compositions containing said protein. The TGF- $\beta$  family of growth factors comprising BMP, TGF, and Inhibin related proteins (Roberts and Sporn, Handbook of Experimental Pharmacology 95 (1990), 419-472) is of particular relevance in a wide range of medical treatments and applications. These factors are useful in processes relating to wound healing and tissue repair. Furthermore, several members of the TGF- $\beta$  family are tissue inductive, especially osteo-inductive, and consequently play a crucial role in inducing cartilage and bone development.

Wozney, Progress in Growth Factor Research 1 (1989), 267-280 and Vale et al., Handbook of Experimental Pharmacology 95 (1990), 211-248 describe different growth factors such as those relating to the BMP (bone morphogenetic proteins) and the Inhibin group. The members of these groups share significant structural similarity. The precursor of the protein is composed of an aminoterminal signal sequence, a propeptide and a carboxyterminal sequence of about 110 amino acids, which is subsequently cleaved from the precursor and represents the mature protein. Furthermore, their members are defined by virtue of amino acid sequence homology. The

004260" 995T0660

mature protein contains the most conserved sequences, especially seven cysteine residues which are conserved among the family members. The TGF- $\beta$ -like proteins are multifunctional, hormonally active growth factors. They also share related biological activities such as chemotactic attraction of cells, promoting cell differentiation and their tissue-inducing capacity, such as cartilage- and bone-inducing capacity. U.S. Patent No. 5,013,649 discloses DNA sequences encoding osteo-inductive proteins termed BMP-2 proteins (bone morphogenetic protein), and U.S. patent applications serial nos. 179 101 and 179 197 disclose the BMP proteins BMP-1 and BMP-3. Furthermore, many cell types are able to synthesize TGF- $\beta$ -like proteins and virtually all cells possess TGF- $\beta$  receptors.

Taken together, these proteins show differences in their structure, leading to considerable variation in their detailed biological function. Furthermore, they are found in a wide variety of different tissues and developmental stages. Consequently, they might possess differences concerning their function in detail, for instance the required cellular physiological environment, their lifespan, their targets, their requirement for accessory factors, and their resistance to degradation. Thus, although numerous proteins exhibiting tissue-inductive, especially osteo-inductive potential are described, their natural role in the organism and, more importantly, their medical relevance must still be elucidated in detail. The occurrence of still-unknown members of the TGF- $\beta$  family relevant for osteogenesis or differentiation/induction of other tissues is strongly suspected. However, a major problem in the isolation of these new TGF- $\beta$ -like proteins is that their functions cannot yet be described precisely enough for the design of a discriminative bioassay. On the other hand, the expected nucleotide sequence homology to known members of the family would be too low to

allow for screening by classical nucleic acid hybridization techniques. Nevertheless, the further isolation and characterization of new TGF- $\beta$ -like proteins is urgently needed in order to get hold of the whole set of induction and differentiation proteins meeting all desired medical requirements. These factors might find useful medical applications in defect healing and treatments of degenerative disorders of bone and/or other tissues like, for example, kidney and liver.

Thus, the technical problem underlying the present invention essentially is to provide DNA sequences coding for new members of the TGF- $\beta$  protein family having mitogenic and/or differentiation-inductive, e.g. osteo-inductive potential.

The solution to the above technical problem is achieved by providing the embodiments characterized in claims 1 to 17. Other features and advantages of the invention will be apparent from the description of the preferred embodiments and the drawings. The sequence listings and drawings will now briefly be described.

SEQ ID NO. 1 shows the nucleotide sequence of MP-52, i.e. the embryo derived sequence corresponding to the mature peptide and most of the sequence coding for the propeptide of MP-52.

Some of the propeptide sequence at the 5'-end of MP-52 has not been characterized so far.

SEQ ID NO. 2 shows the nucleotide sequence of MP-121, i.e. the liver derived sequence corresponding to the mature peptide, the sequence coding for the propeptide of MP-121, and sequences 5' and 3' to the coding region.

0904536-092490

The start codon begins with nucleotide 128 of SEQ ID NO.2. The sequence coding for the mature MP121 polypeptide begins with nucleotide 836 of SEQ ID NO. 2. The stop codon begins with nucleotide 1184 of SEQ ID NO. 2. The sequence coding for the precursor protein has a length of 1056 bp. The sequence coding for the propeptide has a length of 708 bp and the sequence coding for the mature peptide has a length of 348 bp.

SEQ ID NO. 3 shows the amino acid sequence of MP-52 as deduced from SEQ ID NO. 1.

SEQ ID NO. 4 shows the amino acid sequence of MP-121 as deduced from sequence SEQ ID NO.2. The sequence of the mature polypeptide begins with amino acid 237 of SEQ ID NO. 4. The precursor protein has a length of 352 amino acids. The propeptide and the mature peptide have a length of 236 and 116 amino acids, respectively.

SEQ ID NO. 5 shows a part of the nucleotide sequence of the liver derived sequence of MP-121.

SEQ ID NO. 6 shows a part of the nucleotide sequence of the embryo derived sequence of MP-52.

The shorter DNA-sequences SEQ ID NO. 5 and 6 can be useful for example for isolation of further members of the TGF- $\beta$ -protein family.

Figure 1 shows an alignment of the amino acid sequences of MP-52 and MP-121 starting from the first of the seven conserved cysteines with some related proteins. 1a shows the alignment of MP-52 with some members of the BMP protein family; 1b shows the alignment of MP-121 with some members of the Inhibin protein family. \* indicates that the amino acid

is the same in all proteins compared; + indicates that the amino acid is the same in at least one of the proteins compared with MP-52 (Fig. 1a) or MP-121 (Fig. 1b).

Figure 2 shows the nucleotide sequences of the oligonucleotide primer as used in the present invention and an alignment of these sequences with known members of the TGF- $\beta$  family. M means A or C; S means C or G; R means A or G; and K means G or T. 2a depicts the sequence of the primer OD; 2b shows the sequence of the primer OID.

The present invention relates to novel TGF- $\beta$ -like proteins and provides DNA sequences contained in the corresponding genes. Such sequences include nucleotide sequences comprising the sequence

ATGAACTCCATGGACCCCGAGTCCACA and  
CTTCTCAAGGCCAACACAGCTGCAGGCACC

and in particular sequences as illustrated in SEQ ID Nos. 1 and 2, allelic derivatives of said sequences and DNA sequences degenerated as a result of the genetic code for said sequences. They also include DNA sequences hybridizing under stringent conditions with the DNA sequences mentioned above and containing the following amino acid sequences:

Met-Asn-Ser-Met-Asp-Pro-Glu-Ser-Thr or  
Leu-Leu-Lys-Ala-Asn-Thr-Ala-Ala-Gly-Thr.

Although said allelic, degenerate and hybridizing sequences may have structural divergencies due to naturally occurring mutations, such as small deletions or substitutions, they will usually still exhibit essentially the same useful properties, allowing their use in basically the same medical applications.

According to the present invention, the term "hybridization" means conventional hybridization conditions, preferably

664260-95570660

conditions with a salt concentration of 6 x SSC at 62° to 66°C followed by a one-hour wash with 0.6 x SSC, 0.1% SDS at 62° to 66°C. The term "hybridization" preferably refers to stringent hybridization conditions with a salt concentration of 4 x SSC at 62°-66°C followed by a one-hour wash with 0.1 x SSC, 0.1% SDS at 62°-66°C.

Important biological activities of the encoded proteins, preferably MP-52, comprise a mitogenic and osteo-inductive potential and can be determined in assays according to Seyedin et al., PNAS 82 (1985), 2267-2271 or Sampath and Reddi, PNAS 78 (1981), 7599-7603.

The biological properties of the proteins according to the invention, preferably MP-121, may be determined, e.g., by means of the assays according to Wrana et al. (Cell 71, 1003-1014 (1992)), Ling et al. (Proc. Natl. Acad. of Science, 82, 7217-7221 (1985)), Takuwa et al. (Am. J. Physiol., 257, E797-E803 (1989)), Fann and Patterson (Proc. Natl. Acad. of Science, 91, 43-47 (1994)), Broxmeyer et al. (Proc. Natl. Acad. of Science, 85, 9052-9056 (1988)), Green et al. (Cell, 71, 731-739 (1992)), Partridge et al. (Endocrinology, 108, 213-219 (1981)) or Roberts et al. (PNAS 78, 5339-5343 (1981)).

Preferred embodiments of the present invention are DNA sequences as defined above and obtainable from vertebrates, preferably mammals such as pig or cow and from rodents such as rat or mouse, and in particular from primates such as humans.

Particularly preferred embodiments of the present invention are the DNA sequences termed MP-52 and MP-121 which are shown in SEQ ID Nos. 1 and 2. The corresponding transcripts of MP-52 were obtained from embryogenic tissue and code for a

protein showing considerable amino acid homology to the mature part of the BMP-like proteins (see Fig. 1a). The protein sequences of BMP2 (=BMP2A) and BMP4 (=BMP2B) are described in Wozney et al., Science Vol 242, 1528-1534 (1988). The respective sequences of BMP5, BMP6 and BMP7 are described in Celeste et al., Proc.Natl.Acad.Sci. USA Vol 87, 9843-9847 (1990). Some typical sequence homologies, which are specific to known BMP-sequences only, were also found in the propeptide part of MP-52, whereas other parts of the precursor part of MP-52 show marked differences to BMP-precursors. The mRNA of MP-121 was detected in liver tissue, and its correspondig amino acid sequence shows homology to the amino acid sequences of the Inhibin protein chains (see Fig. 1b). cDNA sequences encoding TGF- $\beta$ -like proteins have not yet been isolated from liver tissue, probably due to a low abundance of TGF- $\beta$  specific transcripts in this tissue. In embryogenic tissue, however, sequences encoding known TGF- $\beta$ -like proteins can be found in relative abundance. The inventors have recently detected the presence of a collection of TGF- $\beta$ -like proteins in liver as well. The high background level of clones related to known factors of this group presents the main difficulty in establishing novel TGF- $\beta$ -related sequences from these and probably other tissues. In the present invention, the cloning was carried out according to the method described below. Once the DNA sequence has been cloned, the preparation of host cells capable of producing the TGF- $\beta$ -like proteins and the production of said proteins can be easily accomplished using known recombinant DNA techniques comprising constructing the expression plasmids encoding said protein and transforming a host cell with said expression plasmid, cultivating the transformant in a suitable culture medium, and recovering the product having TGF- $\beta$ -like activity.

09901556-09249



Thus, the invention also relates to recombinant molecules comprising DNA sequences as described above, optionally linked to an expression control sequence. Such vectors may be useful in the production of TGF- $\beta$ -like proteins in stably or transiently transformed cells. Several animal, plant, fungal and bacterial systems may be employed for the transformation and subsequent cultivation process. Preferably, expression vectors which can be used in the invention contain sequences necessary for the replication in the host cell and are autonomously replicable. It is also preferable to use vectors containing selectable marker genes which can be easily selected for transformed cells. The necessary operation is well-known to those skilled in the art.

It is another object of the invention to provide a host cell transformed by an expression plasmid of the invention and capable of producing a protein of the TGF- $\beta$  family. Examples of suitable host cells include various eukaryotic and prokaryotic cells, such as E. coli, insect cells, plant cells, mammalian cells, and fungi such as yeast.

Another object of the present invention is to provide a protein of the TGF- $\beta$  family encoded by the DNA sequences described above and displaying biological features such as tissue-inductive, in particular osteo-inductive and/or mitogenic capacities possibly relevant to therapeutical treatments. The above-mentioned features of the protein might vary depending upon the formation of homodimers or heterodimers. Such structures may prove useful in clinical applications as well. The amino acid sequence of the especially preferred proteins of the TGF- $\beta$ -family (MP-52 and MP-121) are shown in SEQ ID NO. 3 and SEQ ID NO. 4.

It is a further aspect of the invention to provide a process for the production of TGF- $\beta$ -like proteins. Such a process

comprises cultivating a host cell being transformed with a DNA sequence of the present invention in a suitable culture medium and purifying the TGF- $\beta$ -like protein produced. Thus, this process will allow the production of a sufficient amount of the desired protein for use in medical treatments or in applications using cell culture techniques requiring growth factors for their performance. The host cell is obtainable from bacteria such as Bacillus or Escherichia coli, from fungi such as yeast, from plants such as tobacco, potato, or Arabidopsis, and from animals, in particular vertebrate cell lines such as the Mo-, COS- or CHO cell line.

Yet another aspect of the present invention is to provide a particularly sensitive process for the isolation of DNA sequences corresponding to low abundance mRNAs in the tissues of interest. The process of the invention comprises the combination of four different steps. First, the mRNA has to be isolated and used in an amplification reaction using oligonucleotide primers. The sequence of the oligonucleotide primers contains degenerated DNA sequences derived from the amino acid sequence of proteins related to the gene of interest. This step may lead to the amplification of already known members of the gene family of interest, and these undesired sequences would therefore have to be eliminated. This object is achieved by using restriction endonucleases which are known to digest the already-analyzed members of the gene family. After treatment of the amplified DNA population with said restriction endonucleases, the remaining desired DNA sequences are isolated by gel electrophoresis and reamplified in a third step by an amplification reaction, and in a fourth step they are cloned into suitable vectors for sequencing. To increase the sensitivity and efficiency, steps two and three are repeatedly performed, at least two times in one embodiment of this process.

090015E 09249 004260 9510660

In a preferred embodiment, the isolation process described above is used for the isolation of DNA sequences from liver tissue. In a particularly preferred embodiment of the above-described process, one primer used for the PCR experiment is homologous to the polyA tail of the mRNA, whereas the second primer contains a gene-specific sequence. The techniques employed in carrying out the different steps of this process (such as amplification reactions or sequencing techniques) are known to the person skilled in the art and described, for instance, in Sambrook et al., 1989, "Molecular Cloning: A laboratory manual", Cold Spring Harbor Laboratory Press.

It is another object of the present invention to provide pharmaceutical compositions containing a therapeutically-effective amount of a protein of the TGF- $\beta$  family of the present invention. Optionally, such a composition comprises a pharmaceutically acceptable carrier. Such a therapeutic composition can be used in wound healing and tissue repair as well as in the healing of bone, cartilage, or tooth defects, either individually or in conjunction with suitable carriers, and possibly with other related proteins or growth factors. Thus, a therapeutic composition of the invention may include, but is not limited to, the MP-52 encoded protein in conjunction with the MP-121 encoded protein, and optionally with other known biologically-active substances such as EGF (epidermal growth factor) or PDGF (platelet derived growth factor). Another possible clinical application of a TGF- $\beta$ -like protein is the use as a suppressor of the immuno response, which would prevent rejection of organ transplants. The pharmaceutical composition comprising the proteins of the invention can also be used prophylactically, or can be employed in cosmetic plastic surgery. Furthermore, the application of the composition is not limited to humans but can include animals, in particular domestic animals, as well. Possible applications of the pharmaceutical composition

according to the invention include furthermore treatment or prevention of connective tissue, skin, mucous membrane, endothelial, epithelial, neuronal or renal defects, use in the case of dental implants, use as a morphogenic factor used for inducing liver tissue growth, induction of the proliferation of precursor cells or bone marrow cells, for maintaining a differentiated state and the treatment of impaired fertility or for contraception.

Finally, another object of the present invention is an antibody or antibody fragment, which is capable of specifically binding to the proteins of the present invention. Methods to raise such specific antibody are general knowledge. Preferably such an antibody is a monoclonal antibody. Such antibodies or antibody fragments might be useful for diagnostic methods.

The following examples illustrate in detail the invention disclosed, but should not be construed as limiting the invention.

#### Example 1

##### Isolation of MP-121

- 1.1 Total RNA was isolated from human liver tissue (40-year-old-male) by the method of Chirgwin et al., Biochemistry 18 (1979), 5294-5299. Poly A<sup>+</sup> RNA was separated from total RNA by oligo (dT) chromatography according to the instructions of the manufacturer (Stratagene Poly (A) Quick columns).
- 1.2 For the reverse transcription reaction, poly A<sup>+</sup> RNA (1-2.5 µg) derived from liver tissue was heated for 5 minutes to 65°C and cooled rapidly on ice. The reverse transcription reagents containing 27 U RNA guard

(Pharmacia), 2.5  $\mu$ g oligo d(T)<sub>12-18</sub> (Pharmacia) 5 x buffer (250 mM Tris/HCl pH 8.5; 50 mM MgCl<sub>2</sub>; 50 mM DTT; 5 mM each dNTP; 600 mM KCl) and 20 units avian myeloblastosis virus reverse transcriptase (AMV, Boehringer Mannheim) per  $\mu$ g poly (A<sup>+</sup>) RNA were added. The reaction mixture (25  $\mu$ l) was incubated for 2 hours at 42°C. The liver cDNA pool was stored at -20°C.

1.3 The deoxynucleotide primers OD and OID (Fig. 2) designed to prime the amplification reaction were generated on an automated DNA-synthesizer (Biosearch). Purification was done by denaturing polyacrylamide gel electrophoresis and isolation of the main band from the gel by isotachophoresis. The oligonucleotides were designed by aligning the nucleic acid sequences of some known members of the TGF- $\beta$  family and selecting regions of the highest conservation. An alignment of this region is shown in Fig. 2. In order to facilitate cloning, both oligonucleotides contained EcoR I restriction sites and OD additionally contained an Nco I restriction site at its 5' terminus.

1.4 In the polymerase chain reaction, a liver-derived cDNA pool was used as a template in a 50  $\mu$ l reaction mixture. The amplification was performed in 1 x PCR-buffer (16.6 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; 67 mM Tris/HCl pH 8.8; 2 mM MgCl<sub>2</sub>; 6.7  $\mu$ M EDTA; 10 mM  $\beta$ -mercaptoethanol; 170  $\mu$ g/ml BSA (Gibco)), 200  $\mu$ M each dNTP (Pharmacia), 30 pmol each oligonucleotide (OD and OID) and 1.5 units Taq polymerase (AmpliTaq, Perkin Elmer Cetus). The PCR reaction contained cDNA corresponding to 30 ng of poly (A<sup>+</sup>) RNA as starting material. The reaction mixture was overlaid by paraffine and 40 cycles (cycle 1: 80s 93°C/40s 52°C/40s 72°C; cycles 2-9: 60s 93°C/40s 52°C/40s 72°C; cycles 10-29: 60s 93°C/40s 52°C/60s

09901555-09249  
664260-95570660

72°C; cycles 30-31: 60s 93°C/40s 52°C/90s 72°C; cycle 40: 60s 93°C/40s 52°C/420s 72°C) of the PCR were performed. Six PCR-reaction mixtures were pooled, purified by subsequent extractions with equal volumes of phenol, phenol/chloroform (1:1 (v/v)) and chloroform/isoamylalcohol (24:1 (v/v)) and concentrated by ethanol precipitation.

- 1.5 One half of the obtained PCR pool was sufficient for digestion with the restriction enzymes Sph I (Pharmacia) and AlwN I (Biolabs). The second half was digested in a series of reactions by the restriction enzymes Ava I (BRL), AlwN I (Biolabs) and Tfi I (Biolabs). The restriction endonuclease digestions were performed in 100  $\mu$ l at 37°C (except Tfi I at 65°C) using 8 units of each enzyme in a 2- to 12-hour reaction in a buffer recommended by the manufacturer.
- 1.6 Each DNA sample was fractioned by electrophoresis using a 4% agarose gel (3% FMC Nusieve agarose, Biozym and 1% agarose, BRL) in Tris borate buffer (89 mM Trisbase, 89 mM boric acid, 2 mM EDTA, pH 8). After ethidiumbromide staining uncleaved amplification products (about 200 bp; size marker was run in parallel) were excised from the gel and isolated by phenol extraction: an equal volume of phenols was added to the excised agarose, which was minced to small pieces, frozen for 10 minutes, vortexed and centrifuged. The aqueous phase was collected, the interphase reextracted by the same volume TE-buffer, centrifuged and both aqueous phases were combined. DNA was further purified twice by phenol/chloroform and once by chloroform/isoamylalcohol extraction.
- 1.7 After ethanol precipitation, one fourth or one fifth of the isolated DNA was reamplified using the same

664250" 955T.0660

52°C/60s 72°C; cycle 13: 60s 93°C/40s 52°C/420s 72°C). The reamplification products were purified, restricted with the same enzymes as above and the uncleaved products were isolated from agarose gels as mentioned above for the amplification products. The reamplification followed by restriction and gel isolation was repeated once.

- 1.8 After the last isolation from the gel, the amplification products were digested by 4 units EcoR I (Pharmacia) for 2 hours at 37°C using the buffer recommended by the manufacturer. One fourth of the restriction mixture was ligated to the vector pBluescriptII SK+ (Stratagene) which was digested likewise by EcoR I. After ligation, 24 clones from each enzyme combination were further analyzed by sequence analysis. The sample restricted by AlwN I and Sph I contained no new sequences, only BMP6 and Inhibin  $\beta$ A sequences. 19 identical new sequences, which were named MP-121, were found by the Ava I, AlwN I and Tfi I restricted samples. The MP-121 containing plasmids were called pSK MP-121 (OD/OID). One sequence differed from this mainly-found sequence by two nucleotide exchanges. Ligation reaction and transformation in E. coli HB101 were performed as described in Sambrook et al., Molecular cloning: A laboratory manual (1989). Transformants were selected by Ampicillin resistance and the plasmid DNAs were isolated according to standard protocols (Sambrook et al. (1989)). Analysis was done by sequencing the double-stranded plasmids by "dideoxyribonucleotide chain termination sequencing" with the sequencing kit "Sequenase Version 2.0" (United States Biochemical Corporation).

0901556-0949  
64260"9570660

The clone was completed to the 3' end of the c-DNA by a method described in detail by Frohman (Amplifications, published by Perkin-Elmer Corporation, issue 5 (1990), pp 11-15). The same liver mRNA which was used for the isolation of the first fragment of MP-121 was reverse transcribed using a primer consisting of oligo dT (16 residues) linked to an adaptor primer (AGAATTCGCATGCCATGGTCGACGAAGC(T)<sub>16</sub>). Amplification was performed using the adaptor primer (AGAATTCGCATGCCATGGTCGACG) and an internal primer (GGCTACGCCATGAACTTCTGCATA) of the MP-121 sequence. The amplification products were reamplified using a nested internal primer (ACATAGCAGGCATGCCTGGTATTG) of the MP-121 sequence and the adaptor primer. The reamplification products were cloned after restriction with Sph I in the likewise restricted vector pT7/T3 U19 (Pharmacia) and sequenced with the sequencing kit "Sequenase Version 2.0" (United States Biochemical Corporation). Clones were characterized by their sequence overlap to the 3' end of the known MP-121 sequence.

One clone, called p121Lt 3' MP13, was used to isolate a NcoI (blunt ended with T4 polymerase)/SphI fragment. This fragment was ligated into a pSK MP-121 (OD/OID) vector, where the OD primer sequence was located close to the T7 primer sequence of the pSK+ multiple cloning site, opened with SphI/SmaI. The resulting plasmid was called pMP121DFus6. It contains MP-121 specific sequence information starting from position 922 and ending with position 1360 of SEQ ID NO. 2.

- 1.9 Using a DdeI fragment of pMP-121DFus6 as a probe, ranging from nucleotide 931 to nucleotide 1304 of SEQ ID NO. 2, a human liver cDNA library (Clontech, # HL3006b, Lot 36223) was screened by a common method described in



detail by Ausubel et al. (Current Protocols in Molecular Biology, published by Greene Publishing Associates and Wiley-Interscience (1989)). From  $8.1 \times 10^5$  phages, 24 mixed clones were isolated and re-screened using the DdeI fragment. 10 clones were confirmed and the EcoRI fragments subcloned into Bluescript SK (Stratagene, # 212206). EcoRI restriction analysis showed that one clone (SK121 L9.1, deposited by the DSM (#9177) has an insert of about 2.3 kb. This clone contains the complete reading frame of the MP121 gene and further information to the 5' and 3' end in addition to the sequence isolated from mRNA by the described amplification methods. The complete sequence of the EcoRI insert of SK121 L9.1 is shown in SEQ ID NO.2. The reading frame of the MP-121 gene could be confirmed by sequencing of another clone (SK121 L11.1), having the identical reading frame sequence as SK121 L9.1. The beginning of the start codon of the MP-121 sequence of SK121 L9.1 could be determined at position 128 of SEQ ID NO.2, since there are three stop codons in-frame in front of the start codon at positions 62, 77 and 92. The start site of the mature MP-121 is at position 836 of SEQ ID NO.2 in sequence analogy to other members of the TGF- $\beta$ -family, corresponding to amino acid 237 in SEQ ID NO.4. The stop codon is at position 1184 of SEQ ID NO.2.

Plasmid SK121 L9.1 was deposited under number 9177 at DSM (Deutsche Sammlung von Mikroorganismen und Zellkulturen), Mascheroder Weg 1b, Braunschweig, on 26.04.94).

0901556-09249  
064260-9570660

Example 2  
Isolation of MP-52

A further cDNA sequence, MP-52, was isolated according to the above described method (Example 1) by using RNA from human embryo (8-9 weeks old) tissue. The PCR reaction contained cDNA corresponding to 20 ng of poly (A<sup>+</sup>)RNA as starting material. The reamplification step was repeated twice for both enzyme combinations. After ligation, 24 clones from each enzyme combination were further analyzed by sequence analysis. The sample restricted by AlwN I and Sph I yielded a new sequence which was named MP-52. The other clones comprised mainly BMP6 and one BMP7 sequence. The sample restricted by Ava I, AlwN I and Tfi I contained no new sequences, but consisted mainly of BMP7 and a few Inhibin  $\beta$ A sequences.

The clone was completed to the 3' end according to the above described method (Example 1). The same embryo mRNA, which was used for the isolation of the first fragment of MP-52, was reverse transcribed as in Example 1. Amplification was performed using the adaptor primer (AGAATTTCGCATGCCATGGTTCGACG) and an internal primer (CTTGAGTACGAGGCTTTCCACTG) of the MP-52 sequence. The amplification products were reamplified using a nested adaptor primer (ATTCGCATGCCATGGTTCGACGAAG) and a nested internal primer (GGAGCCCACGAATCATGCAGTCA) of the MP-52 sequence. The reamplification products were cloned after restriction with Nco I in a likewise restricted vector (pUC 19 (Pharmacia #27-4951-01) with an altered multiple cloning site containing a unique Nco I restriction site) and sequenced. Clones were characterized by their sequence overlap to the 3' end of the known MP-52 sequence. Some of these clones contain the last 143 basepairs of the 3' end of the sequence shown in SEQ ID NO: 1 and the 0,56 kb 3' non translated region (sequence not shown). One of these was used

as a probe to screen a human genomic library (Stratagene #946203) by a common method described in detail by Ausubel et al. (Current Protocols in Molecular Biology, published by Greene publishing Associates and Wiley-Interscience (1989)). From  $8 \times 10^5$   $\lambda$  phages one phage ( $\lambda$  2.7.4) which was proved to contain an insert of about 20 kb, was isolated and deposited by the DSM (#7387). This clone contains in addition to the sequence isolated from mRNA by the described amplification methods sequence information further to the 5' end. For sequence analysis a Hind III fragment of about 7,5 kb was subcloned in a likewise restricted vector (Bluescript SK, Stratagene #212206). This plasmid, called SKL 52 (H3) MP12, was also deposited by the DSM (# 7353). Sequence information derived from this clone is shown in SEQ ID NO: 1. At nucleotide No. 1050, the determined cDNA and the respective genomic sequence differ by one basepair (cDNA: G; genomic DNA: A). We assume the genomic sequence to be correct, as it was confirmed also by sequencing of the amplified genomic DNA from embryonic tissue which had been used for the mRNA preparation. The genomic DNA contains an intron of about 2 kb between basepairs 332 and 333 of SEQ ID NO: 1. The sequence of the intron is not shown. The correct exon/exon junction was confirmed by sequencing an amplification product derived from cDNA which comprises this region. This sequencing information was obtained by the help of a slightly modified method described in detail by Frohman (Amplifications, published by Perkin-Elmer Corporation, issue 5 (1990), pp 11-15). The same embryo RNA which was used for the isolation of the 3' end of MP-52 was reverse transcribed using an internal primer of the MP-52 sequence oriented in the 5' direction (ACAGCAGGTGGGTGGTGTGGACT). A polyA tail was appended to the 5' end of the first strand cDNA by using terminal transferase. A two step amplification was performed first by application of a primer consisting of oligo dT and an adaptor primer (AGAATTCGCATGCCATGGTCGACGAAGC(T<sub>16</sub>)) and secondly an

09001556 "092499  
654260" 95540660

adaptor primer (AGAATTCGCATGCCATGGTCGACG) and an internal primer (CCAGCAGCCCATCCTTCTCC) of the MP-52 sequence. The amplification products were reamplified using the same adaptor primer and a nested internal primer (TCCAGGGCACTAATGTCAAACACG) of the MP-52 sequence. Consecutively the reamplification products were again reamplified using a nested adaptor primer (ATTCGCATGCCATGGTCGACGAAG) and a nested internal primer (ACTAATGTCAAACACGTACCTCTG) of the MP-52 sequence. The final reamplification products were blunt end cloned in a vector (Bluescript SK, Stratagene #212206) restricted with EcoRV. Clones were characterized by their sequence overlap to the DNA of  $\lambda$  2.7.4.

Plasmid SKL 52 (H3) MP12 was deposited under number 7353 at DSM (Deutsche Sammlung von Mikroorganismen und Zellkulturen), Mascheroder Weg 1b, 3300 Braunschweig, on 10.12.1992.

Phage  $\lambda$  2.7.4. was deposited under number 7387 at DSM on 13.1.1993.

0901556-09249  
664260-95570660

## Claims

1. A DNA sequence encoding a protein of the TGF- $\beta$  family selected from the following group:

(a) a DNA sequence comprising the nucleotides

ATGAACTCCATGGACCCCGAGTCCACA

with the reading frame for the protein starting at the first nucleotide

(b) a DNA sequence comprising the nucleotides

CTTCTCAAGGCCAACACAGCTGCAGGCACC

with the reading frame for the protein starting at the first nucleotide

- (c) DNA sequences which are degenerate as a result of the genetic code to the DNA sequences of (a) and (b)
- (d) allelic derivatives of the DNA sequences of (a) and (b)
- (e) DNA sequences hybridizing to the DNA sequences in (a), (b), (c) or (d) and encoding a protein containing the aminoacid sequence

Met-Asn-Ser-Met-Asp-Pro-Glu-Ser-Thr

or

Leu-Leu-Lys-Ala-Asn-Thr-Ala-Ala-Gly-Thr

- (f) DNA sequences hybridizing to the DNA sequences in (a), (b), (c) and (d) and encoding a protein having essentially the same biological properties.
2. The DNA sequence according to claim 1 which is a vertebrate DNA sequence, a mammalian DNA sequence, preferably a primate, human, porcine, bovine, or rodent DNA sequence, and preferably including a rat and a mouse DNA sequence.
  3. The DNA sequence according to claim 1 or 2 which is a DNA sequence comprising the nucleotides as shown in SEQ ID NO. 1.
  4. The DNA sequence according to claim 1 or 2 which is a DNA sequence comprising the nucleotides as shown in SEQ ID NO. 2.
  5. The DNA sequence according to claim 1 or 2 which is a DNA sequence comprising the nucleotides as shown in SEQ ID NO. 5.
  6. The DNA sequence according to claim 1 or 2 which is a DNA sequence comprising the nucleotides as shown in SEQ ID NO. 6.
  7. A recombinant DNA molecule comprising a DNA sequence according to any one of claims 1 to 6.
  8. The recombinant DNA molecule according to claim 7 in which said DNA sequence is functionally linked to an expression-control sequence.

664260" 955T0660

9. A host containing a recombinant DNA molecule according to claim 7 or 8.
10. The host according to claim 9 which is a bacterium, a fungus, a plant cell or an animal cell.
11. A process for the production of a protein of the TGF- $\beta$  family comprising cultivating a host according to claim 9 or 10 and recovering said TGF- $\beta$  protein from the culture.
12. A protein of the TGF- $\beta$  family encoded by a DNA sequence according to any one of claims 1 to 4 or a fragment thereof encoded by a DNA-sequence according to claim 5 or 6.
13. A protein according to claim 12 comprising the amino acid sequence of SEQ ID NO: 3.
14. A protein according to claim 12 comprising the amino acid sequence of SEQ ID NO. 4.
15. A pharmaceutical composition containing a protein of the TGF- $\beta$  family according to any one of claims 12 to 14, optionally in combination with a pharmaceutically acceptable carrier.
16. The pharmaceutical composition according to claim 15 for the treatment or prevention of bone, cartilage, connective tissue, skin, mucous membrane, endothelial, epithelial, neuronal, renal or tooth defects, for use in the case of dental implants, for use in wound healing or tissue repair processes, as a morphogenic factor used for inducing liver tissue growth, induction of the proliferation of precursor cells or bone marrow cells, for maintaining a differentiated state and for the treatment of impaired fertility or for contraception.

090456 09249  
664260 93540660

17. An antibody or antibody fragment which is capable of specifically binding to a protein of claims 12, 13 or 14.
18. Antibody or antibody fragment according to claim 17 which is a monoclonal antibody.
19. Use of an antibody or antibody fragment according to claims 17 or 18 for diagnostic methods.

000155-09249  
664260-95570660





Abstract

The invention provides DNA sequences encoding novel members of the TGF- $\beta$  family of proteins. The TGF- $\beta$  family comprises proteins which function as growth and/or differentiation factors and which are useful in medical applications. Accordingly, the invention also describes the isolation of the above-mentioned DNA sequences, the expression of the encoded proteins, the production of said proteins and pharmaceutical compositions containing said proteins.

000155 05249  
654260 35510660

a) 1997-1998		b) 1998-1999		c) 1999-2000		d) 2000-2001		e) 2001-2002		f) 2002-2003		g) 2003-2004		h) 2004-2005		i) 2005-2006		j) 2006-2007		k) 2007-2008		l) 2008-2009		m) 2009-2010		n) 2010-2011		o) 2011-2012		p) 2012-2013		q) 2013-2014		r) 2014-2015		s) 2015-2016		t) 2016-2017		u) 2017-2018		v) 2018-2019		w) 2019-2020		x) 2020-2021		y) 2021-2022		z) 2022-2023																																																	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

	10	20	30	40	50	
MP 52	CSRKALHVNF	KDMGWDDWII	APLEYEAFHC	EGLCEFPLRS	HLEPTNHAVI	
BMP 2	CKRHPLYVDF	SDVGWNDWIV	APPGYHAFYC	HGECPFPLAD	HLNSTNHAIV	
BMP 4	CRRHSLYVDF	SDVGWNDWIV	APPGYQAFYC	HGDCPFPLAD	HLNSTNHAIV	
BMP 5	CKKHELYVSF	RDLGWQDWII	APEGYAIFYC	DGECSFPLNA	HMNATNHAIV	
BMP 6	CRKHELYVSF	QDLGWQDWII	APKGYAANYC	DGECSFPLNA	HMNATNHAIV	
BMP 7	CKKHELYVSF	RDLGWQDWII	APEGYAAYYC	EGECAFP LNS	YM NATNHAIV	
	* + * * *	* * * * * +	** * * + *	+ * * * * +	++ * * * *	
	60	70	80	90	100	
MP 52	QTLMNSMDPE	STPPTCCVPT	RLSPISILFI	DSANNVVYKQ	YEDMVVESCG	CR
BMP 2	QTLVNSVNS-	KIPKACCVPT	ELSAISMLYL	DENEKVVLKN	YQDMVVEGCG	CR
BMP 4	QTLVNSVNS-	SIPKACCVPT	ELSAISMLYL	DEYDKVVLKN	YQEMVVEGCG	CR
BMP 5	QTLVHLMFPD	HVPKPCCAPT	KLNAISVLYF	DDSSNVILKK	YRNMVVRSCG	CH
BMP 6	QTLVHLMNPE	YVPKPCCAPT	KLNAISVLYF	DDNSNVILKK	YRNMVVRACG	CH
BMP 7	QTLVHFINPE	TVPKPCCAPT	QLNAISVLYF	DDSSNVILKK	YRNMVVRACG	CH
	*** ++ +	+ * * * * *	* + * * *	* + * + *	* + * * * * *	* +

Fig. 1b

	10	20	30	40
MP121	CCRQEFFVDF	REIGWHDWII	QPEGYAMNFC	IGQCPLHIAG
Inhib $\beta$ A	CCKKQFFVSF	KDIGWNDWII	APSGYHANYC	EGECPSHIAG
Inhib $\beta$ B	CCRQQFFIDF	RLIGWNDWII	APTGYYGNYC	EGSCPAYLAG
Inhib $\alpha$	CHRVALLNISF	QELGWERWIV	YPPSFIFHYC	HGGCGLHIP-
	* + + + +	* + + + +	* + + + +	* + + + +
	50	60	70	80
MP121	MPGIAASFHT	AVLNLLKANT	AAGTTGGGSC	C--VPTARRP
Inhib $\beta$ A	TSGSSLSFHS	TVINHRYMRG	HSPFANLKSC	C--VPTKLRP
Inhib $\beta$ B	VPGSASSFHT	AVVNQYRMRG	LNP-GTVNSC	C--IPTKLST
Inhib $\alpha$	- - - PNL	SLPVPGAPPTPAQP	YSLLPGAQPC	CAALPGTMRP
	+ + + + +	+ + + + +	+ + + + +	+ + + + +
	90	100	110	
MP121	LSLLYYDRDS	NIVKTD-IPD	MVVEACGCS	
Inhib $\beta$ A	MSMLYYDDGQ	NIKKD-IQN	MIVEECGCS	
Inhib $\beta$ B	MSMLYFDDEY	NIVKRD-VPN	MIVEECGCA	
Inhib $\alpha$	LHVRTTSDGG	YSFKYETVPN	LLTQHCACI	
	+ + + + +	+ + + + +	+ + + + +	+ + + + +

## Fig.2a

	Eco RI Nco I
OD	ATGAATTCCCATGGACCTGGGCTGGMAKGAMTGGAT
BMP 2	ACGTGGGGTGGAAATGACTGGAT
BMP 3	ATATTGGCTGGAGTGAATGGAT
BMP 4	ATGTGGGCTGGAATGACTGGAT
BMP 7	ACCTGGGCTGGCAGGACTGGAT
TGF- $\beta$ 1	AGGACCTCGGCTGGAAGTGGAT
TGF- $\beta$ 2	GGGATCTAGGGTGGAAATGGAT
TGF- $\beta$ 3	AGGATCTGGGCTGGAAGTGGGT
inhibin $\alpha$	AGCTGGGCTGGGAACGGTGGAT
inhibin $\beta$ A	ACATCGGCTGGAATGACTGGAT
inhibin $\beta$ B	TCATCGGCTGGAACGACTGGAT

## Fig.2b

	Eco RI
OID	ATGAATTGAGCTGCGTSGGSRACACAGCA
BMP 2	GAGTTCTGTCGGGACACAGCA
BMP 3	CATCTTTTCTGGTACACAGCA
BMP 4	CAGTTCAGTGGGCACACAACA
BMP 7	GAGCTGCGTGGGCGCACAGCA
TGF- $\beta$ 1	CAGCGCCTGCGGCACGCAGCA
TGF- $\beta$ 2	TAAATCTTGGGACACGCAGCA
TGF- $\beta$ 3	CAGGTCCTGGGGCACGCAGCA
inhibin $\alpha$	CCCTGGGAGAGCAGCACAGCA
inhibin $\beta$ A	CAGCTTGGTGGGCACACAGCA
inhibin $\beta$ B	CAGCTTGGTGGGAATGCAGCA

SEQ ID NO. 1

SEQUENCE TYPE: Nucleic Acid  
SEQUENCE LENGTH: 1207 Base Pairs

STRANDEDNESS: Double or Single  
TOPOLOGY: Linear  
MOLECULAR TYPE: DNA or cDNA from mRNA

ORIGINAL SOURCE: -  
ORGANISM: Human  
IMMEDIATE EXPERIMENTAL SOURCE: Embryo Tissue

PROPERTIES: Sequence Coding for Human TGF- $\beta$ -like Protein (MP-52)

ACCGGGCGGC	CCTGAACCCA	AGCCAGGACA	CCCTCCCCAA	ACAAGGCAGG	CTACAGCCCCG	60
GACTGTGACC	CCAAAAGGAC	AGCTTCCCGG	AGGCAAGGCA	CCCCCAAAG	CAGGATCTGT	120
CCCCAGCTCC	TTCCTGCTGA	AGAAGGCCAG	GGAGCCCCGG	CCCCCACGAG	AGCCCAAGGA	180
GCCGTTTCGC	CCACCCCCCA	TCACACCCCCA	CGAGTACATG	CTCTCGCTGT	ACAGGACGCT	240
GTCCGATGCT	GACAGAAAGG	GAGGCAACAG	CAGCGTGAAG	TTGGAGGCTG	GCCTGGCCAA	300
CACCATCACC	AGCTTTATTG	ACAAAGGGCA	AGATGACCGA	GGTCCCGTGG	TCAGGAAGCA	360
GAGGTACGTG	TTTGACATTA	GTGCCCTGGA	GAAGGATGGG	CTGCTGGGGG	CCGAGCTGCG	420
GATCTTGCGG	AAGAAGCCCT	CGGACACGGC	CAAGCCAGCG	GCCCCCGGAG	GCGGGCGGGC	480
TGCCCAGCTG	AAGCTGTCCA	GCTGCCCCAG	CGGCCGGCAG	CCGGCCTCCT	TGCTGGATGT	540
GCGCTCCGTG	CCAGGCCTGG	ACGGATCTGG	CTGGGAGGTG	TTCGACATCT	GGAAGCTCTT	600
CCGAAACTTT	AAGAACTCGG	CCCAGCTGTG	CCTGGAGCTG	GAGGCCTGGG	AACGGGGCAG	660
GGCCGTGGAC	CTCCGTGGCC	TGGGCTTCGA	CCGCGCCGCC	CGGCAGGTCC	ACGAGAAGGC	720
CCTGTTCCTG	GTGTTTGGCC	GCACCAAGAA	ACGGGACCTG	TTCTTTAATG	AGATTAAGGC	780
CCGCTCTGGC	CAGGACGATA	AGACCGTGTA	TGAGTACCTG	TTCAGCCAGC	GGCGAAAACG	840
GCGGGCCCCA	CTGGCCACTC	GCCAGGGCAA	GCGACCCAGC	AAGAACCTTA	AGGCTCGCTG	900
CAGTCGGAAG	GCACTGCATG	TCAACTTCAA	GGACATGGGC	TGGGACGACT	GGATCATCGC	960
ACCCCTTGAG	TACGAGGCTT	TCCACTGCGA	GGGGCTGTGC	GAGTTCCCAT	TGCGTCCCCA	1020
CCTGGAGCCC	ACGAATCATG	CAGTCATCCA	GACCCTGATG	AACTCCATGG	ACCCCGAGTC	1080
CACACCACCC	ACCTGCTGTG	TGCCCACGCG	GCTGAGTCCC	ATCAGCATCC	TCTTCATTGA	1140
CTCTGCCAAC	AACGTGGTGT	ATAAGCAGTA	TGAGGACATG	GTCGTGGAGT	CGTGTGGCTG	1200
CAGGTAG						1207

664260"95T0660

SEQ ID NO.2

SEQUENCE TYPE: Nucleic Acid  
SEQUENCE LENGTH: 2272 Base Pairs

STRANDEDNESS: Double or Single  
TOPOLOGY: Linear  
MOLECULAR TYPE: cDNA from mRNA

ORIGINAL SOURCE: -  
ORGANISM: Human  
IMMEDIATE EXPERIMENTAL SOURCE: Liver Tissue

PROPERTIES: Sequence Coding for Human TGF- $\beta$ -like Protein (MP-121)

CAAGGAGCCA TGCCAGCTGG ACACACACTT CTTCCAGGGC CTCTGGCAGC CAGGACAGAG 60  
TTGAGACCAC AGCTGTTGAG ACCCTGAGCC CTGAGTCTGT ATTGCTCAAG AAGGGCCTTC 120  
CCCAGCAATG ACCTCCTCAT TGCTTCTGGC CTTTCTCCTC CTGGCTCCAA CCACAGTGGC 180  
CACTCCCAGA GCTGGCGGTC AGTGTCCAGC ATGTGGGGGG CCCACCTTGG AACTGGAGAG 240  
CCAGCGGGAG CTGCTTCTTG ATCTGGCCAA GAGAAGCATC TTGGACAAGC TGCACCTCAC 300  
CCAGCGCCCA ACACTGAACC GCCCTGTGTC CAGAGCTGCT TTGAGGACTG CACTGCAGCA 360  
CCTCCACGGG GTCCACAGG GGGCACTTCT AGAGGACAAC AGGGAACAGG AATGTGAAAT 420  
CATCAGCTTT GCTGAGACAG GCCTCTCCAC CATCAACCAG ACTCGTCTTG ATTTTCACTT 480  
CTCCTCTGAT AGAACTGCTG GTGACAGGGA GGTCCAGCAG GCCAGTCTCA TGTTCCTTGT 540  
GCAGCTCCCT TCCAATACCA CTTGGACCTT GAAAGTGAGA GTCCTTGTC TGGGTCCACA 600  
TAATACCAAC CTCACCTTGG CTA CTACAGTA CCTGCTGGAG GTGGATGCCA GTGGCTGGCA 660  
TCAACTCCCC CTAGGGCCTG AAGCTCAAGC TGCCTGCAGC CAGGGGCACC TGACCCTGGA 720  
GCTGGTACTT GAAGGCCAGG TAGCCCAGAG CTCAGTCATC CTGGGTGGAG CTGCCCATAG 780  
GCCTTTTGTG GCAGCCCGGG TGAGAGTTGG GGGCAAACAC CAGATTCACC GACGAGGCAT 840  
CGACTGCCAA GGAGGGTCCA GGATGTGCTG TCGACAAGAG TTTTGTGTGG ACTTCCGTGA 900  
GATTGGCTGG CACGACTGGA TCATCCAGCC TGAGGGCTAC GCCATGAACT TCTGCATAGG 960  
GCAGTGCCCA CTACACATAG CAGGCATGCC TGGTATTGCT GCCTCCTTTC AACTGTCAGT 1020  
GCTCAATCTT CTCAAGGCCA ACACAGCTGC AGGCACCACT GGAGGGGGCT CATGCTGTGT 1080  
ACCCACGGCC CGGCGCCCCC TGTCTCTGCT CTATTATGAC AGGGACAGCA ACATTGTCAA 1140  
GACTGACATA CCTGACATGG TAGTAGAGGC CTGTGGGTGC AGTTAGTCTA TGTGTGGTAT 1200  
GGGCAGCCCA AGGTTGCATG GGAAAACACG CCCCTACAGA AGTGCACCTC CTTGAGAGGA 1260  
GGGAATGACC TCATTCTCTG TCCAGAATGT GGACTCCCTC TTCCTGAGCA TCTTATGGAA 1320  
ATTACCCAC CTTTGA CTTG AAGAAACCTT CATCTAAAGC AAGTCACTGT GCCATCTTCC 1380  
TGACCACTAC CCTCTTTCCT AGGGCATAGT CCATCCCGCT AGTCCATCCC GCTAGCCCCA 1440

CTCCAGGGAC	TCAGACCCAT	CTCCAACCAT	GAGCAATGCC	ATCTGGTTCC	CAGGCAAAGA	1500
CACCCTTAGC	TCACCTTTAA	TAGACCCCAT	AACCCACTAT	GCCTTCCTGT	CCTTTCTACT	1560
CAATGGTCCC	CACTCCAAGA	TGAGTTGACA	CAACCCCTTC	CCCCAATTTT	TGTGGATCTC	1620
CAGAGAGGCC	CTTCTTTTGA	TTCAACCAAAG	TTTAGATCAC	TGCTGCCCAA	AATAGAGGCT	1680
TACCTACCCC	CCTCTTTGTT	GTGAGCCCCT	GTCCCTTCTTA	GTTGTCCAGG	TGAACTACTA	1740
AAGCTCTCTT	TGCATACCTT	CATCCATTTT	TTGTCCCTTCT	CTGCCCTTCT	CTATGCCCTT	1800
AAGGGGTGAC	TTGCCTGAGC	TCTATCACCT	GAGCTCCCCT	GCCCTCTGGC	TTCTTGCTGA	1860
GGTCAGGGCA	TTTCTTATCC	CTGTTCCCTC	TCTGTCTAGG	TGTCATGGTT	CTGTGTAACCT	1920
GTGGCTATTC	TGTGTCCCTA	CACTACCTGG	CTACCCCTT	CCATGGCCCC	AGCTCTGCCT	1980
ACATTCTGAT	TTTTTTTTTT	TTTTTTTTTT	TGAAAAGTTA	AAAATTCCTT	AATTTTTTAT	2040
TCCTGGTACC	ACTACCACAA	TTTACAGGGC	AATATACCTG	ATGTAATGAA	AAGAAAAAGA	2100
AAAAGACAAA	GCTACAACAG	ATAAAAGACC	TCAGGAATGT	ACATCTAATT	GACACTACAT	2160
TGCATTAATC	AATAGCTGCA	CTTTTTTGCA	ACTGTGGCTA	TGACAGTCCT	GAACAAGAAG	2220
GGTTTCCTGT	TTAAGCTGCA	GTAACTTTTC	TGACTATGGA	TCATCGTTCC	TT	2272

[illegible]

•

•

PGGPEPKPGH	PPQTRQATAR	TVTPKGQLPG	GKAPPKAGSV	PSSFLLKKAR	EPGPPPREPKE	60
PFRPPPITPH	EYMLSLYRTL	SDADRKGKNS	SVKLEAGLAN	TITSFIDKGQ	DDRGPVVRKQ	120
RYVFDISALE	KDGLLGAELE	ILRKKPSDTA	KPAAPGGGRA	AQLKLSSCPS	GRQPASLLDV	180
RSVPGLDGSG	WEVFDIWKLF	RNFKNASQLC	LELEAWERGR	AVDLRGLGFD	RAARQVHEKA	240
LFLVFGRTKK	RDLFFNEIKA	RSGQDDKTVY	EYLFSQRRKR	RAPLATRQ GK	RPSKNL KARC	300
SRKALHVNFK	DMGWDDWIIA	PLEYEAFHCE	GLCEFPLRSH	LEPTNHAVIQ	TLMNSMDPES	360
TPPTCCVPTR	LSPISILFID	SANNVVKQY	EDMVVESC GC	R		401



SEQ ID NO. 4

SEQUENCE TYPE: Amino Acid  
SEQUENCE LENGTH: 352 Amino Acids

ORIGINAL SOURCE: -  
ORGANISM: Human

PROPERTIES: Human TGF- $\beta$ -like Protein (MP-121)

MTSSLLLAFL	LLAPTTVATP	RAGGQCPACG	GPTLELESQR	ELLLDLAKRS	ILDKLHLTQR	60
PTLNRPVSRA	ALRTALQHLH	GVPQGALLED	NREQECEIIS	FAETGLSTIN	QTRLDFHFSS	120
DRTAGDREVQ	QASLMFFVQL	PSNTTWTLKV	RVLVLGPHNT	NLTLATQYLL	EVDASGWHQL	180
PLGPEAQ AAC	SQGHLTLELV	LEGQVAQSSV	ILGGAAHRPF	VAARVRVGGK	HQIHRRGIDC	240
QGGSRMCCRQ	EFFVDFREIG	WHDWIIQPEG	YAMNFCIGQC	PLHIAGMPGI	AASFHTAVLN	300
LLKANTAAGT	TGGGSCCVPT	ARRPLSLLYY	DRDSNIVKTD	IPDMVVEACG	CS	352

090455" 09570660

SEQ ID NO. 5

SEQUENCE TYPE: Nucleic Acid  
SEQUENCE LENGTH: 265 Base Pairs

STRANDEDNESS: Double or Single  
TOPOLOGY: Linear  
MOLECULAR TYPE: cDNA from mRNA

ORIGINAL SOURCE: -  
ORGANISM: Human  
IMMEDIATE EXPERIMENTAL SOURCE: Liver Tissue

PROPERTIES: Sequence coding for a Part of the Mature Human TGF- $\beta$ -like Protein  
(MP-121)

CATCCAGCCT GAGGGCTACG CCATGAAC TT CTGCATAGGG CAGTGCCAC TACACATAGC	60
AGGCATGCCT GGTATTGCTG CCTCCTTTCA CACTGCAGTG CTCAATCTTC TCAAGGCCAA	120
CACAGCTGCA GGCACCACTG GAGGGGGCTC ATGCTGTGTA CCCACGGCCC GGCGCCCCCT	180
GTCTCTGCTC TATTATGACA GGGACAGCAA CATTGTCAAG ACTGACATAC CTGACATGGT	240
AGTAGAGGCC TGTGGGTGCA GTTAG	265

090458-0949  
664260-9570660

SEQ ID NO. 6

SEQUENCE TYPE: Nucleic Acid

SEQUENCE LENGTH: 139 Base Pairs

STRANDEDNESS: Double or Single

TOPOLOGY: Linear

MOLECULAR TYPE: cDNA from mRNA

ORIGINAL SOURCE: -

ORGANISM: Human

IMMEDIATE EXPERIMENTAL SOURCE: Embryo Tissue

PROPERTIES: Sequence Coding for a Part of the Mature Human TGF- $\beta$ -like Protein  
(MP-52)

CATCGCACCC CTTGAGTACG AGGCTTTCCA CTGCGAGGGG CTGTGCGAGT TCCCATTGCG	60
CTCCACCTG GAGCCACGA ATCATGCAGT CATCCAGACC CTGATGAACT CCATGGACCC	120
CGAGTCCACA CCACCCACC	139

664260-3570660

Figure 1a

	10	20	30	40	50	
MP 52	CSRKALHVNF	KDMGWDDWII	APLEYEAFHC	EGLCEFPLRS	HLEPTNHAVI	
BMP 2	CKRHPLYVDF	SDVGWNDWIV	APPGYHAFYC	HGECPFPLAD	HLNSTINHAIV	
BMP 4	CRRHSLYVDF	SDVGWNDWIV	APPGYQAFYC	HGDCPFPLAD	HLNSTINHAIV	
BMP 5	CKKHELYVSF	RDLGWQDWII	APEGYAAYYC	DGECFPLNA	HMNATINHAIV	
BMP 6	CRKHELYVSF	QDLGWQDWII	APKGYAANYC	DGECFPLNA	HMNATINHAIV	
BMP 7	CKKHELYVSF	RDLGWQDWII	APEGYAAYYC	EGECAPPLNS	YMNATINHAIV	
	* + * * *	* ** **++ **	* *+ * +* * ***	+ ++ ****		
	60	70	80	90	100	
MP 52	QTLMNSMDPE	STPPTCCVPT	RLSPISILFI	DSANNVVKQ	YEDMVVESC	CR
BMP 2	QTLVNSVNS-	KIPKACCVPT	ELSAISMLYL	DENEKVVLKN	YQDMVVEGC	CR
BMP 4	QTLVNSVNS-	SIPKACCVPT	ELSAISMLYL	DEYDKVVLKN	YQEMVVEGC	CR
BMP 5	QTLVHLMFPD	HVPKPCCAPT	KLNAISVLYF	DDSSNVILKK	YRNMVVRSCG	CH
BMP 6	QTLVHLMNPE	YVPKPCCAPT	KLNAISVLYF	DDNSNVILKK	YRNMVVRACG	CH
BMP 7	QTLVHFINPE	TVPKPCCAPT	QLNAISVLYF	DDSSNVILKK	YRNMVVRACG	CH
	*** +++ ++ + *	***+**	*+ ** *	*	+++ *	* +*****++* *+

000156-0949  
064260-9540660

Figur 1b

	10	20	30	40
MP121	C C R Q E F F V D F	R E I G W H D W I I	Q P E G Y A M N F C	I G Q C P L H I A G
Inhib $\beta$ A	C C K K Q F F V S F	K D I G W N D W I I	A P S G Y H A N Y C	E G E C P S H I A G
Inhib $\beta$ B	C C R Q Q F F I D F	R L I G W N D W I I	A P T G Y Y G N Y C	E G S C P A Y L A G
Inhib $\alpha$	C H R V A L N I S F	Q E L G W E R W I V	Y P P S F I F H Y C	H G G C G L H I P -
	* + + + + + + + *	+ + + + * + * + +	* + + + + * + *	* + + + + + + +
	50	60	70	80
MP121	M P G I A A S F H T	A V L N L L K A N T	A A G T T G G G S C	C - - V P T A R R P
Inhib $\beta$ A	T S G S S L S F H S	T V I N H Y R M R G	H S P F A N L K S C	C - - V P T K L R P
Inhib $\beta$ B	V P G S A S S F H T	A V V N Q Y R M R G	L N P - G T V N S C	C - - I P T K L S T
Inhib $\alpha$	- - - P N L S L P V	P G A P P T P A Q P	Y S L L P G A Q P C	C A A L P G T M R P
	+ + + * + + + + +	+ + + + + + + +	+ + * * + * + +	+ + + + + + + +
	90	100	110	
MP121	L S L L Y Y D R D S	N I V K T D - I P D	M V V E A C G C S	
Inhib $\beta$ A	M S M L Y Y D D G Q	N I I K K D - I Q N	M I V E E C G C S	
Inhib $\beta$ B	M S M L Y F D D E Y	N I V K R D - V P N	M I V E E C G C A	
Inhib $\alpha$	L H V R T T S D G G	Y S F K Y E T V P N	L L T Q H C A C I	
	+ + + + + + + +	+ + + * + + + +	+ + + + * + + +	

Figure 2a

Eco RI Nco I

OD	ATGAATTCCCATGGACCTGGGCTGGMAKGAMTGGAT
BMP 2	ACGTGGGGTGGGAATGACTGGAT
BMP 3	ATATTGGCTGGAGTGAATGGAT
BMP 4	ATGTGGGCTGGGAATGACTGGAT
BMP 7	ACCTGGGCTGGCAGGACTGGAT
TGF- $\beta$ 1	AGGACCTCGGCTGGAAGTGGAT
TGF- $\beta$ 2	GGGATCTAGGGTGGAAATGGAT
TGF- $\beta$ 3	AGGATCTGGGCTGGAAGTGGGT
inhibin $\alpha$	AGCTGGGCTGGGAACGGTGGAT
inhibin $\beta_A$	ACATCGGCTGGAATGACTGGAT
inhibin $\beta_B$	TCATCGGCTGGAACGACTGGAT

Figure 2b

Eco RI

OID	ATGAATTGAGCTGCGTSGGSRACACAGCA
BMP 2	GAGTTCGTGCGGACACAGCA
BMP 3	CATCTTTTCTGGTACACAGCA
BMP 4	CAGTTCAGTGGGCACACAACA
BMP 7	GAGCTGCGTGGGCGCACAGCA
TGF- $\beta$ 1	CAGCGCCTGCGGCACGCAGCA
TGF- $\beta$ 2	TAAATCTTGGGACACGCAGCA
TGF- $\beta$ 3	CAGGTCTTGGGACACGCAGCA
inhibin $\alpha$	CCCTGGGAGAGCAGCACAGCA
inhibin $\beta_A$	CAGCTTGGTGGGCACACAGCA
inhibin $\beta_B$	CAGCTTGGTGGGAATGCAGCA

## Declaration For U.S. Patent Application

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled  
(Insert Title) DNA SEQUENCES ENCODING NOVEL GROWTH/DIFFERENTIATION FACTORS  
the specification of which

- (Check one of blocks 1, 2 or 3. See note A on back of this page)
1. ☐ is attached hereto.
2. ☐ was filed on \_\_\_\_\_ as International PCT Application Serial No. \_\_\_\_\_ and was amended on \_\_\_\_\_ (if applicable)
3. ☒ was filed on August 12, 1994 as U.S. Application Serial No. 08/289,222 and was amended on \_\_\_\_\_ (if applicable)

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claim(s), as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application for which priority is claimed:

	<u>92 102 324.8</u> (Number)	<u>Europe</u> (Country)	<u>12/2/92</u> (Day/Month/Year Filed)	Priority Claimed <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
(List prior foreign applications. See note B on back of this page)	<u>P 44 23 190.3</u> (Number)	<u>DE</u> (Country)	<u>1/7/94</u> (Day/Month/Year Filed)	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
	_____ (Number)	_____ (Country)	_____ (Day/Month/Year Filed)	<input type="checkbox"/> Yes <input type="checkbox"/> No
	_____ (Number)	_____ (Country)	_____ (Day/Month/Year Filed)	<input type="checkbox"/> Yes <input type="checkbox"/> No

(See Note C on back of this page) ☐ See attached list for additional prior foreign applications

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) or PCT International application(s) designating the United States of America listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior application(s) in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT International filing date of this application:

	<u>PCT/EP93/00350</u> (Application Serial No.)	<u>12/2/93</u> (Filing Date)	<u>Pending</u> (Status) (patented, pending, abandoned)
(List prior U.S. Applications or PCT International applications designating the U.S.)	_____ (Application Serial No.)	_____ (Filing Date)	_____ (Status) (patented, pending, abandoned)

And I hereby appoint as principal attorneys David T. Nikaido, Reg. No. 22,663; Charles M. Marmelstein, Reg. No. 25,895; George E. Oram, Jr., Reg. No. 27,931; Robert B. Murray, Reg. No. 22,980; Martin S. Postman, Reg. No. 18,570; E. Marcie Emms, Reg. No. 32,131; Michael G. Gilman, Reg. No. 19,114; Douglas H. Goldhush, Reg. No. 33,125; Kevin C. Brown, Reg. No. 32,402; Monica Chin Kitts, Reg. No. 36,105; Sharon L. Nolan, Reg. No. 36,335; and John R. Fuisz, Reg. No. 37,327.

Please direct all communications to the following address:

NIKAIDO, MARMEIESTEIN, MURRAY & ORAM  
Metropolitan Square  
655 Fifteenth Street, N.W., Suite 330 - G Street Lobby  
Washington, D.C. 20005-5701  
(202) 638-5000 Fax: (202) 638-4810

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

(See Note D on back of this page)

Full name of sole or first inventor Gertrud HOTTEN

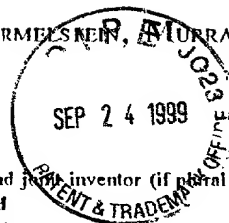
Inventor's signature Gertrud Hotten 11.01.95

Residence Weihwieschenweg 17, Federal Republic of Germany Date

Citizenship German

Post Office Address 69245 Bannwald, Federal Republic of Germany

654260" 955T0660



Full name of second joint inventor, if Helge NEIDHARDT

Inventor's signature Helge Neidhardt 14.1.95

Residence Birkenweg 7, Federal Republic of Germany Date

Citizenship German

Post Office Address 35041 Marburg, Federal Republic of Germany

Full name of third joint inventor, if any Rolf BECHTOLD

Inventor's signature Rolf Bechtold 11.01.95

Residence Carl-Zuckmayer-Str. 21, Federal Republic of Germany Date

Citizenship German

Post Office Address 69126 Heidelberg, Federal Republic of Germany

Full name of fourth joint inventor, if any Jens POHL

Inventor's signature J. Pohl 11.01.95

Residence Kellerswiesen 3, Federal Republic of Germany Date

Citizenship German

Post Office Address 76707 Hambrücken, Federal Republic of Germany

Full name of fifth joint inventor, if any

Inventor's signature

Residence Date

Citizenship

Post Office Address

Full name of sixth joint inventor, if any

Inventor's signature

Residence Date

Citizenship

Post Office Address

Full name of seventh joint inventor, if any

Inventor's signature

Residence Date

Citizenship

Post Office Address

Full name of eighth joint inventor, if any

Inventor's signature

Residence Date

Citizenship

Post Office Address

Full name of ninth joint inventor, if any

Inventor's signature

Residence Date

Citizenship

Post Office Address



SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: HOTTEN, GERTRUD  
NEIDHARDT, HELGE  
BECHTOLD, ROLF  
POHL, JENS

(ii) TITLE OF INVENTION: GROWTH/DIFFERENTIATION FACTORS OF THE TGF-B  
FAMILY

(iii) NUMBER OF SEQUENCES: 53

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: NIKAIDO, MARMELESTEIN, MURRAY & ORAM  
(B) STREET: 655 FIFTEENTH STREET, N. W., G STREET LOBBY,  
SUITE 330  
(C) CITY: WASHINGTON  
(D) STATE: DC  
(E) COUNTRY: USA  
(F) ZIP: 20005-5701

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: UNKNOWN  
(B) FILING DATE: 25-AUG-1999  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/289,222  
(B) FILING DATE: 12-AUG-1994

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: DE P 44 23 190.3  
(B) FILING DATE: 07-JUL-1994

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: EPO 92102324.8  
(B) FILING DATE: 12-FEB-1992

(vii) PRIOR APPLICATION DATA:



# 48  
NB  
M

(A) APPLICATION NUMBER: PCT/EP93/00350  
(B) FILING DATE: 12-FEB-1993

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: KITTS, MONICA CHIN  
(B) REGISTRATION NUMBER: 36,105  
(C) REFERENCE/DOCKET NUMBER: P564-9021

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 202/638-5000  
(B) TELEFAX: 202/638-4810

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1207 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: both  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA or cDNA from mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACCGGGCGGC CCTGAACCCA AGCCAGGACA CCCTCCCCAA ACAAGGCAGG CTACAGCCCG 60  
GACTGTGACC CCAAAAGGAC AGCTTCCCGG AGGCAAGGCA CCCCCAAAAG CAGGATCTGT 120  
GCCCAGCTCC TTCCTGCTGA AGAAGGCCAG GGAGCCCGGG CCCCCACGAG AGCCCAAGGA 180  
GCCGTTTCGC CCACCCCCCA TCACACCCCA CGAGTACATG CTCTCGCTGT ACAGGACGCT 240  
GTCCGATGCT GACAGAAAGG GAGGCAACAG CAGCGTGAAG TTGGAGGCTG GCCTGGCCAA 300  
CACCATCACC AGCTTTATTG ACAAAGGGCA AGATGACCGA GGTCCCGTGG TCAGGAAGCA 360  
GAGGTACGTG TTTGACATTA GTGCCCTGGA GAAGGATGGG CTGCTGGGGG CCGAGCTGCG 420  
GATCTTGCGG AAGAAGCCCT CGGACACGGC CAAGCCAGCG GCCCCCGGAG GCGGGCGGGC 480  
TGCCCAGCTG AAGCTGTCCA GCTGCCCCAG CGGCCGGCAG CCGGCCTCCT TGCTGGATGT 540  
GCGCTCCGTG CCAGGCCTGG ACGGATCTGG CTGGGAGGTG TTCGACATCT GGAAGCTCTT 600

CCGAAACTTT AAGAACTCGG CCCAGCTGTG CCTGGAGCTG GAGGCCTGGG AACGGGGCAG	660
GGCCGTGGAC CTCCGTGGCC TGGGCTTCGA CCGCGCCGCC CGGCAGGTCC ACGAGAAGGC	720
CCTGTTCTTG GTGTTTGGCC GCACCAAGAA ACGGGACCTG TTCTTTAATG AGATTAAGGC	780
CCGCTCTGGC CAGGACGATA AGACCGTGTA TGAGTACCTG TTCAGCCAGC GGCGAAAACG	840
GCGGGCCCCA CTGGCCACTC GCCAGGGCAA GCGACCCAGC AAGAACCTTA AGGCTCGCTG	900
CAGTCGGAAG GCACTGCATG TCAACTTCAA GGACATGGGC TGGGACGACT GGATCATCGC	960
ACCCCTTGAG TACGAGGCTT TCCACTGCGA GGGGCTGTGC GAGTTCCCAT TGCCTCCCA	1020
CCTGGAGCCC ACGAATCATG CAGTCATCCA GACCCTGATG AACTCCATGG ACCCCGAGTC	1080
CACACCACCC ACCTGCTGTG TGCCACGCG GCTGAGTCCC ATCAGCATCC TCTTCATTGA	1140
CTCTGCCAAC AACGTGGTGT ATAAGCAGTA TGAGGACATG GTCGTGGAGT CGTGTGGCTG	1200
CAGGTAG	1207

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2272 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA from mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CAAGGAGCCA TGCCAGCTGG ACACACACTT CTTCCAGGGC CTCTGGCAGC CAGGACAGAG	60
TTGAGACCAC AGCTGTTGAG ACCCTGAGCC CTGAGTCTGT ATTGCTCAAG AAGGGCCTTC	120
CCCAGCAATG ACCTCCTCAT TGCTTCTGGC CTTTCTCCTC CTGGCTCCAA CCACAGTGGC	180
CACTCCCAGA GCTGGCGGTC AGTGTCCAGC ATGTGGGGGG CCCACCTTGG AACTGGAGAG	240
CCAGCGGGAG CTGCTTCTTG ATCTGGCCAA GAGAAGCATC TTGGACAAGC TGCACCTCAC	300

CCAGCGCCCA	AACTGAACC	GCCCTGTGTC	CAGAGCTGCT	TTGAGGACTG	CACTGCAGCA	360
CCTCCACGGG	GTCCCACAGG	GGGCACTTCT	AGAGGACAAC	AGGGAACAGG	AATGTGAAAT	420
CATCAGCTTT	GCTGAGACAG	GCCTCTCCAC	CATCAACCAG	ACTCGTCTTG	ATTTTCACTT	480
CTCCTCTGAT	AGAACTGCTG	GTGACAGGGA	GGTCCAGCAG	GCCAGTCTCA	TGTTCTTTGT	540
GCAGCTCCCT	TCCAATACCA	CTTGGACCTT	GAAAGTGAGA	GTCCTTGTGC	TGGGTCCACA	600
TAATACCAAC	CTCACCTTGG	CTACTCAGTA	CCTGCTGGAG	GTGGATGCCA	GTGGCTGGCA	660
TCAACTCCCC	CTAGGGCCTG	AAGCTCAAGC	TGCCTGCAGC	CAGGGGCACC	TGACCCTGGA	720
GCTGGTACTT	GAAGGCCAGG	TAGCCCAGAG	CTCAGTCATC	CTGGGTGGAG	CTGCCCATAG	780
GCCTTTTGTG	GCAGCCCGGG	TGAGAGTTGG	GGGCAAACAC	CAGATTCACC	GACGAGGCAT	840
CGACTGCCAA	GGAGGGTCCA	GGATGTGCTG	TCGACAAGAG	TTTTTTGTGG	ACTTCCGTGA	900
GATTGGCTGG	CACGACTGGA	TCATCCAGCC	TGAGGGCTAC	GCCATGAACT	TCTGCATAGG	960
GCAGTGCCCA	CTACACATAG	CAGGCATGCC	TGGTATTGCT	GCCTCCTTTC	AACTGCAGT	1020
GCCTCAATCTT	CTCAAGGCCA	ACACAGCTGC	AGGCACCACT	GGAGGGGGCT	CATGCTGTGT	1080
AGCCACGGCC	CGGCGCCCCC	TGTCTCTGCT	CTATTATGAC	AGGGACAGCA	ACATTGTCAA	1140
GACTGACATA	CCTGACATGG	TAGTAGAGGC	CTGTGGGTGC	AGTTAGTCTA	TGTGTGGTAT	1200
GGGCAGCCCA	AGGTTGCATG	GGAAAACACG	CCCCTACAGA	AGTGCACTTC	CTTGAGAGGA	1260
GGGAATGACC	TCATTCTCTG	TCCAGAATGT	GGACTCCCTC	TTCCTGAGCA	TCTTATGGAA	1320
ATTACCCAC	CTTTGACTTG	AAGAAACCTT	CATCTAAAGC	AAGTCACTGT	GCCATCTTCC	1380
TGACCACTAC	CCTCTTTCCT	AGGGCATAGT	CCATCCCGCT	AGTCCATCCC	GCTAGCCCCA	1440
CTCCAGGGAC	TCAGACCCAT	CTCCAACCAT	GAGCAATGCC	ATCTGGTTCC	CAGGCAAAGA	1500
CACCCTTAGC	TCACCTTTAA	TAGACCCCAT	AACCCACTAT	GCCTTCCTGT	CCTTTCTACT	1560
CAATGGTCCC	CACTCCAAGA	TGAGTTGACA	CAACCCCTTC	CCCCAATTTT	TGTGGATCTC	1620
CAGAGAGGCC	CTTCTTTGGA	TTCACCAAAG	TTTAGATCAC	TGCTGCCCAA	AATAGAGGCT	1680
TACCTACCCC	CCTCTTTGTT	GTGAGCCCCT	GTCCTTCTTA	GTTGTCCAGG	TGAACATA	1740

AAGCTCTCTT TGCATACCTT CATCCATTTT TTGTCCTTCT CTGCCTTTCT CTATGCCCTT 1800  
AAGGGGTGAC TTGCCTGAGC TCTATCACCT GAGCTCCCCT GCCCTCTGGC TTCCTGCTGA 1860  
GGTCAGGGCA TTTCTTATCC CTGTTCCCTC TCTGTCTAGG TGTCATGGTT CTGTGTAAC 1920  
GTGGCTATTC TGTGTCCCTA CACTACCTGG CTACCCCCTT CCATGGCCCC AGCTCTGCCT 1980  
ACATTCTGAT TTTTTTTTTT TTTTTTTTTT TGAAAAGTTA AAAATTCCTT AATTTTTTAT 2040  
TCCTGGTACC ACTACCACAA TTTACAGGGC AATATACCTG ATGTAATGAA AAGAAAAAGA 2100  
AAAAGACAAA GCTACAACAG ATAAAAGACC TCAGGAATGT ACATCTAATT GACACTACAT 2160  
TGCATTAATC AATAGCTGCA CTTTTTGCAA ACTGTGGCTA TGACAGTCCT GAACAAGAAG 2220  
GGTTTCCTGT TTAAGCTGCA GTAACCTTTC TGACTATGGA TCATCGTTCC TT 2272

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 401 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Pro	Gly	Gly	Pro	Glu	Pro	Lys	Pro	Gly	His	Pro	Pro	Gln	Thr	Arg	Gln
1				5					10					15	
Ala	Thr	Ala	Arg	Thr	Val	Thr	Pro	Lys	Gly	Gln	Leu	Pro	Gly	Gly	Lys
			20					25					30		
Ala	Pro	Pro	Lys	Ala	Gly	Ser	Val	Pro	Ser	Ser	Phe	Leu	Leu	Lys	Lys
			35				40					45			
Ala	Arg	Glu	Pro	Gly	Pro	Pro	Arg	Glu	Pro	Lys	Glu	Pro	Phe	Arg	Pro
	50					55					60				
Pro	Pro	Ile	Thr	Pro	His	Glu	Tyr	Met	Leu	Ser	Leu	Tyr	Arg	Thr	Leu
65					70					75					80

Ser	Asp	Ala	Asp	Arg 85	Lys	Gly	Gly	Asn	Ser 90	Ser	Val	Lys	Leu	Glu 95	Ala	
Gly	Leu	Ala	Asn 100	Thr	Ile	Thr	Ser	Phe 105	Ile	Asp	Lys	Gly	Gln 110	Asp	Asp	
Arg	Gly	Pro 115	Val	Val	Arg	Lys	Gln 120	Arg	Tyr	Val	Phe	Asp 125	Ile	Ser	Ala	
Leu	Glu 130	Lys	Asp	Gly	Leu	Leu 135	Gly	Ala	Glu	Leu	Arg 140	Ile	Leu	Arg	Lys	
Lys 145	Pro	Ser	Asp	Thr	Ala 150	Lys	Pro	Ala	Ala	Pro 155	Gly	Gly	Gly	Arg 160	Ala	
Ala	Gln	Leu	Lys	Leu 165	Ser	Ser	Cys	Pro	Ser 170	Gly	Arg	Gln	Pro	Ala 175	Ser	
Leu	Leu	Asp	Val 180	Arg	Ser	Val	Pro	Gly 185	Leu	Asp	Gly	Ser	Gly 190	Trp	Glu	
Val	Phe	Asp 195	Ile	Trp	Lys	Leu	Phe 200	Arg	Asn	Phe	Lys	Asn 205	Ser	Ala	Gln	
Leu	Cys 210	Leu	Glu	Leu	Glu	Ala 215	Trp	Glu	Arg	Gly	Arg 220	Ala	Val	Asp	Leu	
Arg 225	Gly	Leu	Gly	Phe	Asp 230	Arg	Ala	Ala	Arg	Gln 235	Val	His	Glu	Lys 240	Ala	
Leu	Phe	Leu	Val	Phe 245	Gly	Arg	Thr	Lys	Lys 250	Arg	Asp	Leu	Phe	Phe 255	Asn	
Glu	Ile	Lys	Ala 260	Arg	Ser	Gly	Gln	Asp 265	Asp	Lys	Thr	Val	Tyr 270	Glu	Tyr	
Leu	Phe	Ser 275	Gln	Arg	Arg	Lys	Arg 280	Arg	Ala	Pro	Leu	Ala 285	Thr	Arg	Gln	
Gly	Lys 290	Arg	Pro	Ser	Lys	Asn 295	Leu	Lys	Ala	Arg	Cys 300	Ser	Arg	Lys	Ala	
Leu 305	His	Val	Asn	Phe	Lys 310	Asp	Met	Gly	Trp	Asp 315	Asp	Trp	Ile	Ile	Ala 320	
Pro	Leu	Glu	Tyr	Glu 325	Ala	Phe	His	Cys	Glu 330	Gly	Leu	Cys	Glu	Phe 335	Pro	

Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Val Ile Gln Thr Leu  
 340 345 350  
 Met Asn Ser Met Asp Pro Glu Ser Thr Pro Pro Thr Cys Cys Val Pro  
 355 360 365  
 Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp Ser Ala Asn Asn  
 370 375 380  
 Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys Gly Cys  
 385 390 395 400  
 Arg

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 352 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Thr Ser Ser Leu Leu Leu Ala Phe Leu Leu Leu Ala Pro Thr Thr  
 1 5 10 15  
 Val Ala Thr Pro Arg Ala Gly Gly Gln Cys Pro Ala Cys Gly Gly Pro  
 20 25 30  
 Thr Leu Glu Leu Glu Ser Gln Arg Glu Leu Leu Leu Asp Leu Ala Lys  
 35 40 45  
 Arg Ser Ile Leu Asp Lys Leu His Leu Thr Gln Arg Pro Thr Leu Asn  
 50 55 60  
 Arg Pro Val Ser Arg Ala Ala Leu Arg Thr Ala Leu Gln His Leu His  
 65 70 75 80  
 Gly Val Pro Gln Gly Ala Leu Leu Glu Asp Asn Arg Glu Gln Glu Cys  
 85 90 95

Glu	Ile	Ile	Ser	Phe	Ala	Glu	Thr	Gly	Leu	Ser	Thr	Ile	Asn	Gln	Thr			
			100					105					110					
Arg	Leu	Asp	Phe	His	Phe	Ser	Ser	Asp	Arg	Thr	Ala	Gly	Asp	Arg	Glu			
		115					120					125						
Val	Gln	Gln	Ala	Ser	Leu	Met	Phe	Phe	Val	Gln	Leu	Pro	Ser	Asn	Thr			
	130					135					140							
Thr	Trp	Thr	Leu	Lys	Val	Arg	Val	Leu	Val	Leu	Gly	Pro	His	Asn	Thr			
145					150					155					160			
Asn	Leu	Thr	Leu	Ala	Thr	Gln	Tyr	Leu	Leu	Glu	Val	Asp	Ala	Ser	Gly			
				165					170					175				
Trp	His	Gln	Leu	Pro	Leu	Gly	Pro	Glu	Ala	Gln	Ala	Ala	Cys	Ser	Gln			
			180					185					190					
Gly	His	Leu	Thr	Leu	Glu	Leu	Val	Leu	Glu	Gly	Gln	Val	Ala	Gln	Ser			
		195					200					205						
Ser	Val	Ile	Leu	Gly	Gly	Ala	Ala	His	Arg	Pro	Phe	Val	Ala	Ala	Arg			
	210					215					220							
Val	Arg	Val	Gly	Gly	Lys	His	Gln	Ile	His	Arg	Arg	Gly	Ile	Asp	Cys			
225					230				235					240				
Gln	Gly	Gly	Ser	Arg	Met	Cys	Cys	Arg	Gln	Glu	Phe	Phe	Val	Asp	Phe			
				245					250					255				
Arg	Glu	Ile	Gly	Trp	His	Asp	Trp	Ile	Ile	Gln	Pro	Glu	Gly	Tyr	Ala			
			260					265					270					
Met	Asn	Phe	Cys	Ile	Gly	Gln	Cys	Pro	Leu	His	Ile	Ala	Gly	Met	Pro			
		275				280						285						
Gly	Ile	Ala	Ala	Ser	Phe	His	Thr	Ala	Val	Leu	Asn	Leu	Leu	Lys	Ala			
	290					295					300							
Asn	Thr	Ala	Ala	Gly	Thr	Thr	Gly	Gly	Gly	Ser	Cys	Cys	Val	Pro	Thr			
305					310					315				320				
Ala	Arg	Arg	Pro	Leu	Ser	Leu	Leu	Tyr	Tyr	Asp	Arg	Asp	Ser	Asn	Ile			
				325				330						335				
Val	Lys	Thr	Asp	Ile	Pro	Asp	Met	Val	Val	Glu	Ala	Cys	Gly	Cys	Ser			
			340					345					350					



(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 265 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA from mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CATCCAGCCT GAGGGCTACG CCATGAACTT CTGCATAGGG CAGTGCCAC TACACATAGC 60  
AGGCATGCCT GGTATTGCTG CCTCCTTTCA CACTGCAGTG CTCAATCTTC TCAAGGCCAA 120  
CACAGCTGCA GGCACCACTG GAGGGGGCTC ATGCTGTGTA CCCACGGCCC GGCGCCCCCT 180  
GCTCTGCTC TATTATGACA GGGACAGCAA CATTGTCAAG ACTGACATAC CTGACATGGT 240  
AGTAGAGGCC TGTGGGTGCA GTTAG 265

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 139 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA from mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CATCGCACCC CTTGAGTACG AGGCTTTCCA CTGCGAGGGG CTGTGCGAGT TCCCATTGCG 60  
CTCCCACCTG GAGCCCACGA ATCATGCAGT CATCCAGACC CTGATGAACT CCATGGACCC 120  
CGAGTCCACA CCACCCACC 139

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATGAACTCCA TGGACCCCGA GTCCACA

27

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GTCTCAAGG CCAACACAGC TGCAGGCACC

30

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Asn Ser Met Asp Pro Glu Ser Thr  
1 5

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Leu Leu Lys Ala Asn Thr Ala Ala Gly Thr  
1 5 10

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AGAATTCGCA TGCCATGGTC GACGAAGCTT TTTTTTTTTT TTTTTT

46

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGAATTCGCA TGCCATGGTC GACG

24

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGCTACGCCA TGAACCTCTG CATA

24

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ACATAGCAGG CATGCCTGGT ATTG

24

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTTGAGTACG AGGCTTTCCA CTG

23

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATTTCGCATGC CATGGTCGAC GAAG

24

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGAGCCCACG AATCATGCAG TCA

23

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ACAGCAGGTG GGTGGTGTGG ACT

23

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCAGCAGCCC ATCCTTCTCC

20

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TCCAGGGCAC TAATGTCAAA CACG

24

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: both  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ACTAATGTCA AACACGTACC TCTG

24

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 102 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Cys Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met Gly Trp Asp  
1 5 10 15  
Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys Glu Gly  
20 25 30  
Leu Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala  
35 40 45  
Val Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr Pro Pro  
50 55 60  
Thr Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile  
65 70 75 80  
Asp Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val  
85 90 95  
Glu Ser Cys Gly Cys Arg  
100

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Cys Lys Arg His Pro Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn  
1 5 10 15  
Asp Trp Ile Val Ala Pro Pro Gly Tyr His Ala Phe Tyr Cys His Gly  
20 25 30  
Glu Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala  
35 40 45  
Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys Ala  
50 55 60  
Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp  
65 70 75 80  
Glu Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp Met Val Val Glu  
85 90 95  
Gly Cys Gly Cys Arg  
100

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Cys Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn





Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val  
85 90 95

Arg Ser Cys Gly Cys His  
100

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 102 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Cys Arg Lys His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Gln  
1 5 10 15  
Asp Trp Ile Ile Ala Pro Lys Gly Tyr Ala Ala Asn Tyr Cys Asp Gly  
20 25 30  
Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala  
35 40 45  
Ile Val Gln Thr Leu Val His Leu Met Asn Pro Glu Tyr Val Pro Lys  
50 55 60  
Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe  
65 70 75 80  
Asp Asp Asn Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val  
85 90 95  
Arg Ala Cys Gly Cys His  
100

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 102 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear



His Thr Ala Val Leu Asn Leu Leu Lys Ala Asn Thr Ala Ala Gly Thr  
50 55 60

Thr Gly Gly Gly Ser Cys Cys Val Pro Thr Ala Arg Arg Pro Leu Ser  
65 70 75 80

Leu Leu Tyr Tyr Asp Arg Asp Ser Asn Ile Val Lys Thr Asp Ile Pro  
85 90 95

Asp Met Val Val Glu Ala Cys Gly Cys Ser  
100 105

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 106 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Cys Cys Lys Lys Gln Phe Phe Val Ser Phe Lys Asp Ile Gly Trp Asn  
1 5 10 15

Asp Trp Ile Ile Ala Pro Ser Gly Tyr His Ala Asn Tyr Cys Glu Gly  
20 25 30

Glu Cys Pro Ser His Ile Ala Gly Thr Ser Gly Ser Ser Leu Ser Phe  
35 40 45

His Ser Thr Val Ile Asn His Tyr Arg Met Arg Gly His Ser Pro Phe  
50 55 60

Ala Asn Leu Lys Ser Cys Cys Val Pro Thr Lys Leu Arg Pro Met Ser  
65 70 75 80

Met Leu Tyr Tyr Asp Asp Gly Gln Asn Ile Ile Lys Lys Asp Ile Gln  
85 90 95

Asn Met Ile Val Glu Glu Cys Gly Cys Ser  
100 105

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 105 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Cys	Cys	Arg	Gln	Gln	Phe	Phe	Ile	Asp	Phe	Arg	Leu	Ile	Gly	Trp	Asn
1				5				10					15		
Asp	Trp	Ile	Ile	Ala	Pro	Thr	Gly	Tyr	Tyr	Gly	Asn	Tyr	Cys	Glu	Gly
			20					25					30		
Ser	Cys	Pro	Ala	Tyr	Leu	Ala	Gly	Val	Pro	Gly	Ser	Ala	Ser	Ser	Phe
		35					40					45			
His	Thr	Ala	Val	Val	Asn	Gln	Tyr	Arg	Met	Arg	Gly	Leu	Asn	Pro	Gly
	50					55					60				
Thr	Val	Asn	Ser	Cys	Cys	Ile	Pro	Thr	Lys	Leu	Ser	Thr	Met	Ser	Met
65					70					75					80
Leu	Tyr	Phe	Asp	Asp	Glu	Tyr	Asn	Ile	Val	Lys	Arg	Asp	Val	Pro	Asn
			85						90					95	
Met	Ile	Val	Glu	Glu	Cys	Gly	Cys	Ala							
			100					105							

(xii) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 105 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Cys	His	Arg	Val	Ala	Leu	Asn	Ile	Ser	Phe	Gln	Glu	Leu	Gly	Trp	Glu
1				5				10					15		

Arg	Trp	Ile	Val	Tyr	Pro	Pro	Ser	Phe	Ile	Phe	His	Tyr	Cys	His	Gly
			20					25					30		
Gly	Cys	Gly	Leu	His	Ile	Pro	Pro	Asn	Leu	Ser	Leu	Pro	Val	Pro	Gly
		35					40					45			
Ala	Pro	Pro	Thr	Pro	Ala	Gln	Pro	Tyr	Ser	Leu	Leu	Pro	Gly	Ala	Gln
	50					55					60				
Pro	Cys	Cys	Ala	Ala	Leu	Pro	Gly	Thr	Met	Arg	Pro	Leu	His	Val	Arg
65					70					75					80
Thr	Thr	Ser	Asp	Gly	Gly	Tyr	Ser	Phe	Lys	Tyr	Glu	Thr	Val	Pro	Asn
				85					90					95	
Leu	Leu	Thr	Gln	His	Cys	Ala	Cys	Ile							
			100					105							

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 36 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATGAATTCCC ATGGACCTGG GCTGGMAKGA MTGGAT

36

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ACGTGGGGTG GAATGACTGG AT

22

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

ATATTGGCTG GAGTGAATGG AT

22

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

ATGTGGGCTG GAATGACTGG AT

22

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

ACCTGGGCTG GCAGGACTGG AT

22

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 22 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

AGGACCTCGG CTGGAAGTGG AT

22

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 22 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GGGATCTAGG GTGGAAATGG AT

22

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 22 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

AGGATCTGGG CTGGAAGTGG GT

22

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 22 base pairs  
    (B) TYPE: nucleic acid



(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

AGCTGGGCTG GGAACGGTGG AT

22

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

ACATCGGCTG GAATGACTGG AT

22

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TCATCGGCTG GAACGACTGG AT

22

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

ATGAATTCGA GCTGCGTSGG SRCACAGCA

29

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GAGTTCTGTC GGGACACAGC A

21

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CATCTTTTCT GGTACACAGC A

21

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

CAGTTCAGTG GGCACACAAC A

21

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GAGCTGCGTG GGCGCACAGC A

21

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

CAGCGCCTGC GGCACGCAGC A

21

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

TAAATCTTGG GACACGCAGC A

21

(2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

CAGGTCCTGG GGCACGCAGC A

21

(2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CCCTGGGAGA GCAGCACAGC A

21

(2) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

CAGCTTGGTG GGCACACAGC A

21

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CAGCTTGGTG GGAATGCAGC A

21

664260"995T0660